From Substance Use to Homelessness or Vice Versa?

McVicar, D., Moschion, J., & van Ours, J. C. (2015). From Substance Use to Homelessness or Vice Versa? *Social Science & Medicine*, *136-137*, 89-98. https://doi.org/10.1016/j.socscimed.2015.05.005

From Substance Use to Homelessness or Vice Versa?*

Duncan McVicar
Queen's University Management School, Queen's University Belfast

Julie Moschion (Corresponding author)
Melbourne Institute of Applied Economic and Social Research, University of Melbourne;
EconomiX, University of Nanterre; moschion@unimelb.edu.au.

Jan C. van Ours
Department of Economics and CentER, Tilburg University; Department of Economics,
University of Melbourne; CEPR (London); CESifo (Munich); IZA (Bonn).

*Acknowledgements: This paper uses unit record data from Journeys Home: Longitudinal Study of Factors Affecting Housing Stability (Journeys Home). The study was initiated and is funded by the Australian Government Department of Social Services (DSS). The Department of Employment has provided information for use in Journeys Home and it is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to DSS, the Department of Employment or the Melbourne Institute. The authors would also like to thank Dan O'Flaherty, Dave Ribar, Steve Stillman, the two anonymous referees and seminar participants at Humboldt University and the University of Melbourne for useful comments on earlier drafts.

Abstract

Homelessness is associated with substance use, but whether substance use precedes or follows homelessness is unclear. We investigate the nature of the relationship between homelessness and substance use using data from the unique Australian panel dataset Journeys Home collected in 4 surveys over the period from October 2011 to May 2013. Our data refer to 1325 individuals who were homeless or at risk of becoming homeless. We investigate dynamics in homelessness and substance use over the survey period. We find that the two are closely related: homeless individuals are more likely to be substance users and substance users are more likely to be homeless. These relationships, however, are predominantly driven by observed and unobserved individual characteristics which cause individuals to be both more likely to be homeless and to be substance users. Once we take these personal characteristics into account it seems that homelessness does not affect substance use, although we cannot rule out that alcohol use increases the probability that an individual becomes homeless. These overall relationships also hide some interesting heterogeneity by 'type' of homelessness.

1 Introduction

The prevalence of homelessness is difficult to measure, but a recent Department of Housing and Urban Development snapshot estimate for the United States (US) suggested around 630,000 people were sleeping on the streets or in shelters in January 2012 (HUD, 2012). This estimate would be considerably higher were those 'doubling up' with family or friends or in other forms of insecure housing included, and higher still if the number in question referred to people experiencing homelessness within a period of time rather than at a single point in time. For example, Link et al. (1994) estimated that 4.6% of the US population had been homeless at some point over the 5 years between 1986 and 1990. More recently O'Flaherty (2012) has again stressed the importance of thinking about homelessness from a dynamic perspective.

Not only does homelessness deprive individuals and families of a basic human need (Curtis et al., 2013), it is also strongly associated with a wide range of other social problems. For example, levels of substance (ab)use are far higher among the homeless than among the wider population (e.g. Greene et al., 1997; Shinn et al, 1998; Early, 2005; Kemp et al., 2006). Indeed there is a widely-held view that homelessness and substance use are self-reinforcing, i.e. that substance use *causes* homelessness (e.g. Allgood and Warren, 2003; Early, 2005), and that homelessness *causes* substance use (e.g. Shinn et al., 1998; Johnson and Chamberlain, 2008), or both (e.g. Johnson et al., 1997; Neale, 2001). Whether a causal link exists between homelessness and substance use, and if so of what magnitude and in which direction(s), are obviously crucial questions for policy makers and service providers designing and delivering interventions in this area.

Unfortunately the lack of consensus in this regard reflects an evidence base which is at best patchy. At the heart of this problem lies a dearth of suitable data on the substance use of representative samples of individuals experiencing homelessness and individuals at risk of

homelessness observed over time. This is not an easy population to reach, let alone reach repeatedly. As a result many studies are based on small-scale, ad hoc, cross-section surveys (Scutella and Johnson, 2012). These surveys also tend to be of very specific – often acutely homeless or acutely using – populations such as clients of treatment centres, other service providers, or 'skid row' communities in urban centres (e.g. Teeson et al., 2000; Booth et al., 2002). This is also the case, albeit to a lesser extent, for larger cross-sections studied in this literature like the National Survey of Homeless Assistance Providers and Clients (NSHAPC) (see e.g. Early 2005). Other studies exploit cross-sections enhanced with retrospective information (including the NSHAPC), but again these tend to be of specific provider-based populations (e.g. Allgood and Warren, 2003; Johnson and Chamberlain, 2008). Studies using longitudinal data also tend to be based on small samples of similarly specific populations, often with a very limited time dimension, and with little attempt made to deal with unobservable confounders (e.g. Allgood et al., 1997; Zlotnick et al., 2003; Rice et al., 2005; Kemp et al., 2006). The net result is a body of literature from which it is difficult to draw general conclusions and from which we can learn little regarding causality even within the study populations themselves.

Two more promising studies using longitudinal data, covering homeless and at-risk-of-homelessness individuals, are Shinn et al. (1998) and Fertig and Reingold (2008). Fertig and Reingold (2008) exploit data from the Fragile Families and Child Wellbeing Study tracking around 5000 low-income parents from the birth of their child for a further three years. Once they control for a host of observable individual and contextual factors they find no evidence that mothers with a 'drug problem' – those who report that their drinking or drug use interfered with their work or personal relationships over the previous 12 months – are any more likely to be homeless at the one-year follow up interview. Shinn et al. (1998) use survey data on around 250 families requesting shelter accommodation in New York City and a

similar number of comparison families drawn from welfare records, with both groups re-interviewed five years later. They also find little evidence of a substance abuse impact on homelessness once other observable factors are controlled. Neither study, however, examines whether homelessness impacts on substance use in a multivariate model, both draw on only small numbers of people experiencing homelessness (from quite particular populations), and both have limited information on substance use (single dummies in each case, with the bar set high to be counted as a substance abuser).

The current paper is the first to examine the dynamic inter-relationships between homelessness and substance use, potentially running in both directions, in a large-scale, broadly–based longitudinal survey, drawing on a population covering individuals experiencing differing degrees of homelessness and comparable individuals not currently homeless but at risk of homelessness, and with richly detailed data on substance use. The data come from the Australian Journeys Home (JH) study, which is unique in its scale, detail and coverage (see Scutella et al., 2012), and has not previously been used to study the links between homelessness and substance use beyond a handful of descriptive tables and brief accompanying discussions in the in-house research reports linked to the release of the first four waves of data (Scutella et al., 2012; Chigavazira et al., 2013; Johnson et al., 2013; Scutella et al., 2013).

Specifically, we address four research questions. First, what are the extent, nature and persistence of substance use among homeless people and those at risk of homelessness in the JH sample? Second, what is the association between substance use and homelessness status in the JH sample? Third, to what extent might these associations be driven by causal relationships from substance use to homelessness, from homelessness to substance use, or both? Fourth, do these relationships vary according to type of homelessness?

In addressing these questions we make a number of specific contributions. We provide the most detailed description of the use of alcohol, cannabis, and other illegal/street drugs among JH respondents during the first 2 years (4 waves) of the survey to date. Second, we demonstrate strong point-in-time associations between substance use and homelessness for this broad-based sample. Third, we show that these associations are predominantly driven by observed and, crucially, unobserved individual characteristics which cause individuals to be both substance users and homeless. Once we take these characteristics into account, and appealing to the arrow of time to infer the direction of any remaining relationship, we conclude that homelessness does not seem to affect substance use in the next 6 months, while only risky alcohol use seems to increase the probability that an individual becomes homeless in the next 6 months. Finally, we show that some substance use behaviours appear to impact heterogeneously on different types of homelessness. The implication is that conclusions regarding the associations between substance use and homelessness are likely to be sensitive to the definition of homelessness used, in particular whether those 'doubling up' with friends and family and those in other temporary accommodation are included along with those sleeping rough.

The remainder of the paper is set out as follows. Section 2 provides descriptive information about homelessness and substance use in the JH sample. Section 3 presents our exploratory analysis of the determinants of homelessness and section 4 the determinants of substance use. Section 5 discusses sensitivity analysis and presents estimates by homelessness type. Section 6 concludes. An online appendix includes further data details and results.

2 Homelessness and substance use

2.1 The JH data and variable definitions

JH is a longitudinal dataset with information on a sample of recipients of any income support (i.e. welfare) payment who are either homeless or at-risk of homelessness in Australia (Scutella et al., 2012). The Melbourne Institute ran this survey for which data collection was approved by the Human Ethics Committee of the University of Melbourne. Here we use the first four waves collected between September 2011 and May 2013 focussing on the balanced panel, i.e. the 1325 respondents who were interviewed in all four waves of the survey. Despite the disadvantaged nature of the target population, both the response rate at wave 1 (61.9%) and the retention rate in the balanced panel at wave 4 (79% of wave 1 respondents) were high. Note the resulting similarity in characteristics between the full wave 1 and balanced panel samples (see Table 1). Further, by controlling for time-invariant unobserved heterogeneity and for a wide range of time-varying characteristics, the fixed-effects regression framework adopted in this paper helps to alleviate any remaining concerns relating to non-random response. More information on the data collection is provided in Appendix A. [INSERT LINK TO ONLINE FILE A]

Table 1 around here

Homelessness can be defined in different ways and with different thresholds. Here we follow Johnson and Chamberlain (2008) in adopting the so-called 'cultural definition' of homelessness used by the Australian Bureau of Statistics in their efforts to enumerate the homeless population in Australia. This definition distinguishes three 'types' of homelessness. Primary homelessness is defined here as sleeping rough or squatting in abandoned buildings. This is essentially the acute unsheltered homelessness of Curtis et al. (2013) or the street

homelessness of Early (2004). Secondary homelessness is defined as staying with relatives or friends temporarily with no alternative, i.e. the 'doubling-up' of Fertig and Reingold (2008) and Curtis et al. (2013). Tertiary homelessness is defined as staying in a caravan, boarding house, hotel or crisis accommodation, a category which includes but is not limited to the acute sheltered homeless of Curtis et al. (2013). Together these three categories constitute a broad conceptualization of homelessness very similar to that under the 2009 Homeless Emergency Assistance and Rapid Transition to Housing Act, and similar to that used by Link et al. (1994) and Curtis et al. (2013).

Specifically, we construct dummies for primary homeless, secondary homeless, tertiary homeless, and their union, for each of the four waves. In wave 1 the dummies are equal to 1 if the individual has been primary/secondary/tertiary homeless in the last 6 months and 0 otherwise. In waves 2-4 the dummies are equal to 1 if the individual has been primary/secondary/tertiary homeless since the last interview. Note that although it seems reasonable to interpret primary homelessness as a more extreme form of homelessness than either secondary or tertiary homelessness, there is no such intuitive ordering between the secondary and tertiary homelessness categories.

Similarly, substance use can be defined in different ways and with different thresholds. Many studies specify a single generic dummy for a substance use *disorder* – sometimes but not always reflecting data limitations – whether based on self-reports of whether one's own drinking or drug use interferes or has interfered with work or personal relationships (Fertig and Reingold, 2008), interviewer/case worker reports as to whether the individual is receiving or has received treatment for substance dependency or whether they meet diagnostic criteria for alcohol or drug dependency (e.g. Allgood et al., 1997; Shinn et al., 1998; Early, 2005; Johnson and Chamberlain, 2008), or some closely related information (e.g. Johnson et al., 1997). Other studies define separate dummies along these lines for alcohol and

drug disorders, and interestingly, their interaction (e.g. Booth et al., 2002; Allgood and Warren, 2003; Zlotnick et al., 2003). By setting a high bar to be counted as a 'problem substance user', and not distinguishing between different levels of substance use below this threshold, these studies are focussing attention very much on the right hand tail of the substance use intensity distribution. Definitions based on receipt of treatment also confound substance use with individual and institutional *responses* to substance use.

In contrast, here we sidestep the question of whether an individual's substance use represents a disorder, and define variables based purely on (self-reported) use and frequency/intensity of use, separately identifying use of alcohol, cannabis, and illegal/street drugs other than cannabis. Specifically, we construct dummies for whether respondents drank alcohol at risky levels (defined, following the Australian National Health and Medical Research Council definition, as 21 or more standard drinks per week on average) since the last wave, whether they used cannabis since the last wave and whether they used illegal/street drugs other than cannabis since the last wave. These are lower bars than those used in the majority of studies cited above, and it seems likely that many of those classed as not having a substance use disorder in these earlier studies would be classed as substance users here. Direct comparisons with these existing studies are therefore unlikely to be particularly informative. Having said that, when we test sensitivity to varying the thresholds for the substance use dummies our conclusions remain unchanged. We also consider tobacco use, treating it as a placebo variable in our model for homelessness, using a dummy variable for whether respondents have smoked tobacco daily during the last six months/since the previous wave.

2.2 Substance use, homelessness, and their association in the JH sample

Table 1 shows the prevalence of substance use in each wave for the balanced panel. Around two thirds of the sample report daily use of tobacco in each wave, followed (in order of prevalence) by cannabis use, risky drinking and use of other illegal street drugs (reported by between 10% and 15% in each wave). Overall, outcomes (and observed characteristics) generally 'improve' over time, e.g. rates of homelessness fall from around two thirds in wave 1 to just under one half in wave 4. Given that we restrict our attention to the balanced panel this does not reflect attrition but the combined effects of time and time since selection into the sample. Specifically, because those flagged as homeless were initially oversampled, and because homelessness is for many a transitory state, we would expect the prevalence of homelessness (and substance use given it is correlated with homelessness) to fall over time in the balanced panel. In all models that follow we include wave dummies to control for this and for common time effects.

Table 2 adds detail on lifetime substance use and on persistence of use across survey waves. Comparing to the general Australian population these descriptive statistics, together with those in Table 1, confirm the relatively high prevalence of substance use among JH respondents. With the exception of tobacco use, however, these behaviours are often transitory, with few engaged in risky drinking, cannabis or other illegal street drug use in all four waves.

Table 2 about here

To capture intensity of use, Figure 1 shows the distribution of average consumption by substance over the survey period for substance users. Most cannabis users used only a few days per month but there is a long right tail with six percent of users using every day at every wave. Most users of other illegal street drugs only used on a few days per month. To

illustrate the intensity of risky drinking we use information on binge drinking, which is defined as drinking five or more standard drinks on any one occasion. Most of those who report binge drinking report doing so on few occasions during the preceding month, but again there is a substantial minority reporting frequent binge drinking. Few respondents smoked more than 25 cigarettes per day on average over the four waves, but around two thirds of smokers smoked at least ten cigarettes per day.

Figure 1 about here

Table 3 gives a further indication of the positive association between homelessness and substance use, with those reporting being homeless at some time during the survey period also reporting higher prevalence of substance use across all behaviours with the exception of injecting street drugs. Table 3 also shows heterogeneity in the strength of the association between homelessness and substance use by *type* of homelessness, with the percentage of respondents having used a particular substance over the course of JH always higher among those who have experienced primary homelessness over the course of JH than among those who experienced secondary or tertiary but not primary homelessness.

Table 3 about here

In the previous section we stressed the importance of analysing homelessness from a dynamic perspective. Table 4 provides summary information for the JH sample by comparing homelessness experiences between survey waves (6 months intervals). For many, homelessness (and equally, not being homeless) persists across waves. But for others homelessness, or a particular type of homelessness, is a temporary state. Note that secondary or tertiary homelessness is more often a step on the way from primary homelessness to not being homeless, and less often a step on the way in the other direction.

Table 4 about here

3 Determinants of homelessness

In the previous section we presented evidence of a number of associations between homelessness and substance use. However, these associations are point-in-time and unconditional, i.e. they ignore dynamics and they may be driven by differences in individual and contextual characteristics that influence both homelessness and substance use. To further investigate the association between substance use and homelessness we estimate a series of linear models for homelessness, initially separately for each substance use behaviour with no controls other than wave dummies, then adding observable time invariant and time varying controls, then replacing the time-invariant controls with individual fixed effects. In each case substance use is included both contemporaneously and lagged one wave. Our final (and preferred) model includes all four substance use behaviours in a single equation – note that information on injection of drugs was only collected in waves 3 and 4 so we omit it from our regression analysis – with individual fixed effects and time-varying observable controls, as given by (1):

$$H_{it} = \beta_1 T_{it} + \beta_2 T_{it-1} + \beta_3 A_{it} + \beta_4 A_{it-1} + \beta_5 C_{it} + \beta_6 C_{it-1} + \beta_7 D_{it} + \beta_8 D_{it-1} + \beta_9 X_{it} + \gamma_i + \tau_t + \eta_{it} \,, \quad (1)$$

where H_{it} denotes homelessness (whether primary, secondary or tertiary) of individual i at time t; T, A, C and D denote tobacco use, risky alcohol use (21+ units per week), cannabis use and illegal/street drug use respectively; X_{it} represents observable controls (listed and defined in Table 1); γ_i and τ_t are individual and time fixed effects respectively and η_{it} is the error term.

In estimating (1) we are investigating whether substance use over the survey period precedes homelessness, controlling for all time-invariant differences between individuals,

whether observed or unobserved, and for observable time-varying differences. Conditional on these controls, we interpret a statistically significant relationship from lagged substance use to current homelessness as indicating a potentially causal relationship from substance use to homelessness, appealing to the arrow of time to rule out causation in the opposite direction. Johnson and Chamberlain (2008) discuss possible mechanisms for such a relationship, including breakdown of family relationships and financial strain resulting from substance use. We do not place any causal interpretation on significant contemporaneous associations between homelessness and substance use; although neither can we rule out that a causal relationship may in part underlie any such associations.

Table 5 shows the resulting estimates of β_1 through β_8 and associated standard errors. Column 1 shows positive and statistically significant associations between homelessness and contemporaneous tobacco, alcohol, cannabis and illegal/street drug use, and positive and statistically significant associations with lagged alcohol, cannabis, and illegal/street drug use, but not lagged tobacco use. For example, those reporting illegal/street drug use in any given wave are 7.7 percentage points more likely to be homeless in that wave, and are 9.3 percentage points more likely to be homeless in the following wave. Associations between homelessness and current and lagged cannabis use, and also lagged risky alcohol use, are of similar magnitude.

Table 5 around here

Conditioning on observables substantially improves the fit of the model in each case but kills almost all of the statistically significant correlations with contemporaneous substance use (primarily because the coefficients fall in magnitude). The associations between homelessness and lagged alcohol, cannabis and illegal/street drug use also fall in magnitude, but remain statistically significant at the 95%, 90% and 90% levels respectively. The explanation for this

lies in the observable confounders, which take signs largely as we would expect and in line with earlier studies where variables are shared (e.g. Early, 2004). For example, homelessness in the JH sample is negatively associated with having dependent children and positively associated with being male, being indigenous, with dummies for having experienced physical or sexual violence in the last 6 months, and with having all/most friends being homeless, all of which are also correlated with substance use. There are also some variables that one might expect to be statistically significant that here are not, including employment experience over the last 6 months (negative but not statistically significant), the amount of outstanding debt (positive but not statistically significant) and various dummies for parental/caregiver behaviour, including alcohol and drug use (positive and statistically insignificant for male caregivers but negative and statistically insignificant for female caregivers), when the individual was aged 14 years. The full set of estimates for these covariates is reported in Table S1 in the online appendix [INSERT LINK TO ONLINE FILE B].

Although we have a rich set of observable controls, there may of course be unobservable factors that continue to confound the estimated relationships between the lagged substance use variables and homelessness. Conditioning on individual fixed effects washes out these unobserved factors to the extent that they are time-invariant. This has a mixed impact on the estimated substance use coefficients, increasing the magnitudes of the lagged and contemporaneous risky alcohol use and illegal/street drugs estimates and decreasing (and rendering statistically insignificant) the magnitude of the contemporaneous tobacco use estimate.

Because these substance use behaviours are themselves correlated with one another, however, we cannot be sure whether the remaining significant associations with lagged substance use in column 3 of Table 5 are being driven by risky drinking, illegal drug use, or both. Our preferred fixed effects model therefore includes all eight substance use variables

together and the relevant estimates are given in the final column of Table 5. In this model only lagged risky alcohol use remains statistically significant, and we interpret this, somewhat tentatively, as *potentially* indicating a causal effect. The magnitude of this estimated coefficient suggests that risky alcohol use increases the probability of homelessness during the next six months by 10.7 percentage points. Contemporaneous risky alcohol use is also associated with homelessness although we hazard no particular causal interpretation on this. The estimated impact of lagged illegal/street drug use other than cannabis on homelessness is less than half the size and nowhere near statistically significant. The estimated coefficient on lagged cannabis use is just below the 90% statistical significance threshold, but again is less than half the magnitude of the lagged risky drinking coefficient. We interpret these estimates, again somewhat tentatively, as indicating a lack of any causal effect large enough and precisely estimated enough to show up convincingly in our data.

Although one can easily envisage causal mechanisms from alcohol, cannabis and other illegal/street drug use to homelessness, such mechanisms seem intuitively unlikely in the case of tobacco use. The estimated coefficient on lagged tobacco use can therefore arguably be interpreted as something akin to a falsification test, where a positive and statistically significant coefficient would suggest the observable controls and individual fixed effects fail to adequately wash out selection effects. Table 5 shows this estimated coefficient to be statistically insignificant and negative. Given positive correlations between tobacco use and the other substance use variables (Scutella et al., 2013), this supports our interpretation of the risky drinking effect as potentially causal.

4 Determinants of substance use

In the same way we explore potential causal effects of homelessness on substance use, estimating linear models separately for each substance, first including only wave dummies as controls, then including observable time-varying and time-invariant controls as in (1), and finally replacing the time-invariant observed controls with individual fixed effects. Our preferred model – the fixed effects model – is given by (2):

$$S_{it} = \alpha_1 H_{it} + \alpha_2 H_{it-1} + \alpha_3 Z_{it} + \theta_i + \varphi_t + \mu_{it}, \quad (2)$$

where S_{it} denotes substance use at time t, H_{it} and H_{it-1} are dummies for whether respondent i was homeless since the last interview or between the interview in t-2 and the interview in t-1, respectively, Z_{it} is a vector of controls which overlaps very closely with X_{it} (the 'homeless friends' variable is replaced by a 'using friends' variable), θ_i and φ_t are individual and time fixed effects and μ_{it} is the error term. As in Section 3, if a parameter of lagged homelessness has a significant effect on substance use, we interpret this as indicating a potentially causal effect of homelessness on substance use. Adapting to a subculture of substance use among the homeless and/or using substances as a coping mechanism are both possible mechanisms for such a causal effect (Johnson and Chamberlain, 2008). Estimates are presented in Table 6.

Table 6 about here

The first column of Table 6 shows positive and statistically significant associations between homelessness at time *t* and alcohol, cannabis and other illegal/street drug use at time *t*, consistent with the associations presented in Table 5. There are also positive and statistically significant associations between lagged homelessness and current tobacco and current cannabis use. Most of these associations do not disappear when observable controls are included (column 2), although this substantially increases the model R^2s. In this case

significant controls – which again take intuitive signs – include age (positive, and age square negative for all four substances use behaviours), male (positive for all four substance use behaviours), indigenous status (positive for risky alcohol use), the level of education (negative for tobacco use, risky alcohol use and cannabis use), being born in an English-speaking country (positive for tobacco use, risky alcohol use and cannabis use), having dependent children (negative for risky alcohol use, cannabis use and other illegal drug use), having experienced physical violence in the last 6 months (positive for all four behaviours), having experienced sexual violence in the last 6 months (positive for other illegal drug use), having friends who are mostly drug users (positive for all four behaviours), the amount of outstanding debt (positive for cannabis use and other illegal drug use), having a male caregiver at age 14 with drug or alcohol problems (positive for tobacco and alcohol use), and reporting having experienced emotional abuse or neglect as a child (positive for cannabis use and other illegal drug use). Table S2 reports the full set of estimates for these covariates [INSERT LINK TO ONLINE FILE B].

The key step in (2), however, is the inclusion of individual fixed effects. This kills all remaining positive associations between lagged homelessness and current substance use: coefficients on lagged homelessness (and for that matter current homelessness) in the models for tobacco use, cannabis use and other illegal drug use are small in magnitude and nowhere near statistical significance at standard levels. The only effect of homelessness that remains statistically significant is for risky alcohol use, but it is *negative*, small in magnitude and only significant at the 90 percent level. We interpret these estimates as suggesting a lack of any substantial causal effect from homelessness as defined here to substance use as defined here.

5 Sensitivity Analysis and Extensions

First consider sensitivity to the precise definitions of our substance use variables. We explore three specific changes: varying the threshold at which alcohol consumption is defined as risky, replacing the cannabis use variable with a dummy for daily cannabis use, and replacing the other illegal/street drug variable with a dummy for weekly use. Lowering or raising the threshold at which alcohol consumption is defined as risky does not impact qualitatively on our conclusion of a possible causal effect on homelessness. Setting the dummy equal to 1 for 15+ drinks per week and equal to zero otherwise – which results in an additional 6% or so of the sample switching from not risky drinking to risky drinking in each wave – halves the magnitude of the coefficient on lagged risky drinking in (1), but it remains statistically significant at the 90% level. When we raise the threshold to 28+ drinks per week, the relevant coefficient falls in between the low and medium threshold coefficients, with risky alcohol use increasing the probability of homelessness during the next six months by 8 percentage points (statistically significant at the 95% level). Varying the risky drinking definition in (2) suggests that the negative effect from homelessness to risky drinking is larger when the risky drinking threshold is lower, becoming statistically insignificant when the threshold is set at 28+ drinks per week. Tightening the frequency of use thresholds for cannabis use and other illegal/street drug use also has no effect on our conclusion of zero impact from these behaviours on homelessness in (1) and zero impact of homelessness on these behaviours in (2). Results are presented in Table S3 in the online appendix [INSERT LINK TO ONLINE FILE B].

Second, we replace time-varying controls observed at time t with their lags at t-1, in both (1) and (2). In doing so we are primarily concerned with whether any of these 'controls', e.g. divorce/separation between t-1 and t, might in fact be capturing causal mechanisms through which substance use impacts on homelessness or vice versa. The obvious trade-off is a lesser degree of control for current time-varying factors. Table S4 [INSERT LINK TO ONLINE

FILE B] shows that lagging the controls in this way has no impact on our conclusions of a positive effect of risky alcohol use on homelessness in (1), no other substance use impacts on homelessness in (1), and zero effects from homelessness to substance use in (2).

Third, we investigate whether our main results vary by gender (they do not) and whether they are sensitive to including a lagged dependent variable in (1) and (2) and estimating by GMM (they are not). Results are presented in Tables S5 and S6 [INSERT LINK TO ONLINE FILE B]. To explore the possible impact of non-random attrition we re-estimate our preferred models on the unbalanced panel including a forward-looking dummy for absence from the next wave. The attrition dummy is statistically insignificant in all but one case (being primary homeless at time t marginally increases the probability of attrition in the next wave), and conclusions regarding the relationships between substance use and homelessness are in all cases unaffected (see Table S7 [INSERT LINK TO ONLINE FILE B]). We explore various other potential sources of sensitivity, both to the estimated coefficients and the standard errors, including restricting the sample to be common across all the different specifications of (1) and (2) and clustering at 'sample cluster' level instead of individual level. None of our conclusions are affected by these changes.

Next consider disaggregating the homelessness dummy by type of homelessness. Table 7 presents estimates from our preferred specification of (1) but with the homelessness dummy first replaced by a primary homelessness dummy (equal to 1 if the individual has been primary homeless since the last wave and 0 otherwise), and second by an ordered categorical variable equal to 2 for those having experienced primary homelessness since the last wave, equal to 1 for those having experienced on secondary or tertiary homelessness since the last wave, and equal to 0 otherwise. Both models are estimated linearly.

Table 7 around here

The right hand column of Table 7 just confirms our earlier conclusions: there is no statistically significant impact from lagged tobacco, cannabis or other illegal/street drug use on homelessness, and there is a potential causal impact from lagged risky alcohol use, whether homelessness is measured as a single dummy or as an ordered categorical variable.

Replacing the homelessness dummy with a primary homelessness dummy, however, does point to some interesting heterogeneity. This means we are now comparing those primary homeless with those not homeless and those secondary/tertiary homeless pooled together. We find a smaller effect of lagged risky drinking in this case, although it remains statistically significant at the 95% level. The suggestion is that risky drinking impacts most strongly on secondary and/or tertiary homelessness. Studies using a broader definition of homelessness might therefore be more likely to find evidence of alcohol use impacts than those using a narrower definition. There is also a somewhat surprising result in the primary homelessness version of (1): the coefficient on lagged illegal/street drug use becomes negative and statistically significant. Our explanation for this is that drug use impacts positively on secondary/tertiary homelessness – this is what we pick up in the positive but statistically insignificant coefficient in Table 5 – rather than that it impacts negatively on primary homelessness. Restricting the sample to those who are either primary homeless or not homeless at time t gives an insignificant coefficient of -0.15. Excluding the primary homeless and replacing the primary homelessness dummy with a secondary/tertiary homelessness dummy gives a positive but insignificant coefficient of 0.56. Evidently, our tentative conclusion of no causal impact from illegal/street drug use on homelessness hides some interesting heterogeneity, where small negative impacts on primary homelessness offset small positive impacts on secondary and tertiary homelessness. Again, the definition of homelessness appears to be important.

Table 8 presents estimates from the fixed effects version of (2) with the current and lagged homelessness dummies disaggregated into separate dummies for primary, secondary and tertiary homelessness. This makes absolutely no difference to our conclusion of zero impact from homelessness on substance use: none of the 12 lagged homelessness dummies across the four substance use models is statistically significant.

Table 8 around here

6 Summary and Conclusions

This paper exploits unique longitudinal data for a large and broadly-based sample of homeless and at-risk-of-homelessness individuals to examine the dynamics of homelessness, substance use and, using standard panel data methods, the associations between substance use and homelessness. In doing so it makes a number of significant contributions to a mostly descriptive literature bedevilled by a dearth of suitable data to examine these issues in anything other than small and often very specialised samples. First, the paper provides a detailed analysis of the use of tobacco, alcohol, cannabis, and other illegal/street drugs among survey respondents over a period of two years. Second, the paper demonstrates strong cross-sectional associations between substance use and homelessness among this sample. Third, and most importantly, the paper exploits the arrow of time in fixed effect regression models to show that these associations between homelessness and substance use, in both directions, are predominantly driven by observed and unobserved factors which cause individuals to be both substance users and homeless. Once these factors are taken into account it appears that homelessness does not affect substance use, while only risky alcohol use impacts on homelessness. Fourth, we highlight some interesting heterogeneity underlying these overall

conclusions, including that risky alcohol use appears to impact more heavily on secondary and tertiary homelessness than on primary homelessness. These conclusions are highly robust.

Direct comparisons with existing studies are difficult given differences in the nature of the data, in the definitions of key variables and in modelling approach. Nevertheless in concluding that there may be a causal link from risky alcohol use to homelessness in the JH sample we are broadly in line with some earlier studies for the US (e.g. Allgood and Warren, 2003; Early, 2005) but not others (e.g. Shinn et al., 1998; Fertig and Reingold, 2008). Similarly, in demonstrating a lack of homelessness effects on substance use we contrast with Johnson et al. (1997), Shinn et al. (1998) and Johnson and Chamberlain (2008), although only Johnson et al. (1997) based their conclusions in this regard on a multivariate model for substance use (and one with few controls at that). The heterogeneity by homelessness type demonstrated here also implies that conclusions regarding the associations between substance use and homelessness elsewhere in the literature may be sensitive to the definition of homelessness used. Recent changes in the way homelessness is defined by the US Administration make this finding particularly timely.

With all the usual caveats about the extent to which these conclusions will generalise across contexts, we draw out three tentative implications for policy. First, high levels of substance use among the homeless may be sufficient grounds to target substance use treatment programs at homeless populations, even in the absence of any causal relationship between substance use and homelessness. Second, the lack of a causal relationship from homelessness to substance use suggests more support for interventions to target reduced substance use among existing users rather than interventions trying to prevent take up of substances by those yet to become users. Third, the possible existence of an economically significant causal relationship from some forms of substance use to homelessness seemingly

offers the potential for early interventions on alcohol use among at risk populations to help reduce entry into homelessness.

References

Allgood, S., Moore, M., & Warren Jr., R. S. (1997). The duration of sheltered homelessness in a small city. *Journal of Housing Economics*, 6, 60–80.

Allgood, S., & Warren Jr., R. S. (2003). The duration of homelessness: evidence from a national survey. *Journal of Housing Economics*, 12, 273–290.

Booth, B., Sullivan, G., Koegel, P., & Burnam, A. (2002). Vulnerability factors for homelessness associated with substance dependence in a community sample of homeless adults. *The American Journal of Drug and Alcohol Abuse*, 28(3), 429-452.

Chigavazira, A., Johnson, G., Moschion, J., Scutella, R., Tseng, Y. and Wooden M. (2013). Journeys Home *Research Report 2*. Report prepared for the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs.

Curtis, M. A., Corman, H., Noonan, K. & Reichman. N. E. (2013). Life shocks and homelessness. *Demography*, 50, 2227–2253.

Early, D. W. (2004). The determinants of homelessness and the targeting of housing assistance. *Journal of Urban Economics*, 55, 195–214.

Early, D. (2005). An empirical investigation of the determinants of street homelessness. *Journal of Housing Economics*, 14, 27–47.

Fertig, A. R., & Reingold, D. A. (2008). The characteristics and causes of homelessness among at risk families with children in twenty American cities. *Social Service Review*, 82, 485–510.

Greene, J.M., Ennett, S.T., & Ringwalt, C.L. (1997). Substance use among runaway and homeless youth in three national samples. *American Journal of Public Health*, 87, 229-235.

HUD (2012). *The 2012 Point-in-Time Estimates of Homelessness*, Volume I of the 2012 Annual Homeless Assessment Report. The U.S. Department of Housing and Urban Development.

Johnson, G., & Chamberlain, C. (2008). Homelessness and substance abuse: which comes first? *Australian Social Work*, 61(4), 342-356.

Johnson, T., Freels, S., Parsons, J. & Vangeest, J. (1997). Substance abuse and homelessness: social selection or social adaptation. *Addiction*, 92(4), 437-445.

Johnson, G., Moschion, J., Scutella, R., Tseng, Y. and Wooden M. (2013). Journeys Home *Research Report 3*. Report prepared for the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs.

Kemp, P., Neale, J., & Robertson, M. (2006). Homelessness among problems drug users: prevalence, risk factors and triggers events. *Health and Social Care in the Community*, 14, 319-328.

Link, B. G., Susser, E., Stueve, A., Phelan, J., Moore, R. E., & Struening, E. (1994). Lifetime and five-year prevalence of homelessness in the United States. *American Journal of Public Health*, 84, 1907–1912.

Melbourne Institute (2013). Journeys Home Wave 4 Technical Report – Fieldwork, Response and Weighting. Report prepared for the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs.

Neale, J. (2001). Homelessness amongst drug users: a double jeopardy explored. *The International Journal of Drug Policy*, 12, 353-369.

O'Flaherty, B. (2012). Individual homelessness: entries, exits, and policy. *Journal of Housing Economics*, 21, 77-100.

Rice, E., Milburn, N.G., Rotheram-Borus, M.J., Mallett, S., & Rosenthal, D. (2005). The effects of peer group network properties on drug use among homeless youth. *American Behavioural Scientist*, 48(8), 1102-1123.

Scutella, R., & Johnson, G. (2012). Locating and designing 'Journeys Home': a literature review. Melbourne Institute Working Paper No. 11/12, University of Melbourne.

Scutella, R., Johnson, G., Moschion, J., Tseng, Y. and Wooden M. (2012). Journeys Home *Research Report 1*. Report prepared for the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs.

Scutella, R., Chigavazira, A., Killackey, E., Herault, N., Johnson, G., Moschion, J. (2013). Journeys Home *Research Report 4*. Report prepared for the Australian Government Department of Social Services.

Shinn, M., Weitzman, B.C., Stojanovic, D., Knickman, J.R., Jimenez, L., Duchon, & L., Krantz, D.H. (1998). Predictors of homelessness among families in New York City: from shelter request to housing stability. *American Journal of Public Health*, 88, 1651–1657.

Teesson, M., Hodder, T., & Buhrich, N. (2000). Substance use disorders among homeless people in Inner Sydney. *Social Psychiatry and Psychiatric epidemiology*, 35, 451-456.

Wooden, M., Bevitt, A., Chigavazira, A., Greer, N., Johnson, G., Killackey, E., Moschion, J., Scutella, R., Tseng, Y. and Watson, N. (2012). Introducing "Journeys Home". *Australian Economic Review*, 45(3), 368-378.

Zlotnick, C., Tam, T., & Robertson, M.J. (2003). Disaffiliation, substance use, and exiting homelessness. *Substance Use and Misuse*, 38, 577–599.

Table 1: Sample Characteristics

	Wave 1 respondents	Balanced panel W1 charact.	Balanced panel W2 charact.	Balanced panel W3 charact.	Balanced panel W4 charact.
Time-varying variables					
Homeless in last 6m	66.6	67.3	56.7	53.3	49.1
Risky drinking (21+/wk) in last 6m	18.8	17.3*	17.0	14.8	15.5
Used cannabis in last 6m	37.3	38.7	34.9	37.4	32.7
Used illegal/street drugs in last 6m	14.4	14.2	9.9	14.5	10.2
Age, years	32.0	31.7	32.2	32.7	33.2
Smoked daily in last 6m	68.0	68.2	67.4	67.8	66.8
Divorced / separated	19.6	19.9	19.2	19.5	19.5
Experienced physical violence in last 6m	18.4	18.9	17.8	17.7	15.2
Experienced sexual violence in last 6m	2.1	2.0	1.7	1.8	1.6
Employed in last 6m	21.5	22.1	27.3	29.3	27.4
Proportion of time employed in last 6m	0.9	0.9	0.8	0.7	0.7
Contacts with family less than once a month	20.6	19.4	19.3	19.1	18.9
Total outstanding debt (in AUD)	5744.0	5696.4	5206.5	5879.4	5029.7
All/most friends are homeless	9.5	9.2	6.9	6.8	5.5
All/most friends are using illegal drugs	18.6	19.0	16.3	17.1	13.7
Completed Year 12	28.6	29.7	32.0	33.8	35.8
Holds a tertiary qualification	11.4	12.6	12.4	12.3	11.6
Lives with dependent children	17.1	17.5*	20.7	21.5	23.5
Time-invariant variables		*Balanced panel W1-W4 characteristics*			
Male	59.1	58.7*			
Indigenous (including Torres Straight Islander)	21.6	19.4*			
Born in an English speaking country	91.8	91.9			
Spent some time in State care	25.1	23.6			
Parents were divorced/separated at age 14	32.2	33.5			
Parents were dead at age 14	6.5	6.7			
Conflicts with parents at age 14	7.1	6.7			
Experienced emotional abuse, physical or sexual violence as a child	65.9	68.8*			
Male caregiver					
Had an alcohol or drug problem	28.5	28.7			
Spent time in jail	10.3	10.6			
Spent time in hospital because mental health pbs	5.1	4.6			
Was unemployed more than 6 m	16.5	14.7*			
Had a gambling problem	8.4	8.3			
Female caregiver					
Had an alcohol or drug problem	17.2	17.0			
Spent time in jail	2.0	2.3			
Spent time in hospital because mental health pbs	10.7	11.2			

Was unemployed more than 6 m	38.1	37.8
Had a gambling problem	7.3	6.5
Number of observations	1,682	1,325

Note: Weighted using the wave 1 weights for the first column and the balanced panel weights for columns 2 to 5. * denote significant differences at the 95% level between the characteristics of the balanced panel (at wave 1) and those of attriters (at wave 1).

Table 2: Prevalence and persistence of substance use (%)

	Alcohol - 21+ standard drinks/wk	Cannabis	Illegal/ Street drugs	Injecting illegal/ street drugs	Tobacco - daily use
Ever tried	-	79.6	52.2	23.4	-
Ever used on a regular basis(1)	-	49.4	22.8	-	84.1
Ever used over survey period (2 years)	31.0	52.4	25.3	8.6	76.6
1 wave only	11.3	10.6	11.8	5.2	3.8
2 waves	6.9	9.7	5.4	3.2	6.0
3 waves	5.3	12.4	5.1	-	10.7
All 4 waves	4.5	18.3	2.5	-	54.8
Australian population(2)	20.1	14.7		0.4	15.1

Notes: Balanced panel (1,325 observations). Injection data only available for waves 3 & 4.
(1) Daily for cannabis and weekly for street drugs.
(2) These figures are from the Australian Institute of Health and Welfare (AIHW) (2011b) 2010 National Drug Strategy Household Survey and give use for the last 12 months. The AIHW (2011b) defines risk levels of alcohol consumption according to the recently revised Australian Alcohol Guidelines: "For healthy men and women, drinking no more than 2 standard drinks on any day reduces the lifetime risk of harm from alcohol-related disease or injury". Respondents having on average 21 or more standard drinks per week are considered at risk. Note the 20 percent figure for alcohol in the bottom row is the proportion of the Australian population who report drinking more than 2 standard drinks *on days they are drinking* (which may not be every day). In the JH sample, the equivalent figure is between 55 percent and 59 percent of respondents at each wave.

Table 3: Percentage of respondents using substances by homeless experience during Journeys Home (%)

	Alcohol - 21+ standard drinks/wk	Cannabis	Illegal/ Street drugs	Injecting illegal/ street drugs	Tobacco - daily use	N
Never homeless	25.3	41.0	17.5	8.5	69.3	220
Ever secondary or tertiary	27.4	50.3	23.4	5.6	75.3	804
Ever primary	44.1	67.7	36.9	17.1	85.5	290

Note: Balanced panel (1,325 observations). The "never homeless" were "never homeless" during JH; the "ever secondary or tertiary" have been secondary or tertiary homeless but never primary; the "ever primary" have been primary homeless during JH (11 respondents' homeless status is not observed at every wave and therefore cannot be classified).

Table 4 Transitions in the degree of homelessness

	Degree of homelessness time t				
Degree of homelessness time t+1	Primary	Secondary or tertiary	Not homeless	Unknown	Share (%)
Primary	48	3	2	9	6
Secondary or tertiary	32	69	21	25	46
Not homeless	17	27	76	43	46
Unknown	4	1	1	23	1
Total (%)	100	100	100	100	100
Total (number)	391	1954	1592	37	3,975
Share (%)	10	49	40	1	100

Note: the figures show the homelessness situation of respondents between survey waves which are 6 months apart, averaged over the 4 waves of the balanced panel (and thus over 3*1325 = 3975 observations).

Table 5: Determinants of homelessness at time t

	No controls	Observable controls	Fixed effects	FE, all subst. together
Alcohol - 21+ standard drinks/wk				
In t	0.043*	0.022	0.083***	0.071**
	(0.026)	(0.026)	(0.032)	(0.032)
In t-1	0.080***	0.054**	0.108***	0.107***
	(0.025)	(0.024)	(0.032)	(0.033)
N	3,765	3,765	3,765	3,732
R^2 / within R^2	0.018	0.121	0.026	0.028
Cannabis				
In t	0.073***	0.030	0.017	0.010
	(0.022)	(0.022)	(0.027)	(0.029)
In t-1	0.072***	0.039*	0.045*	0.040
	(0.021)	(0.021)	(0.025)	(0.027)
N	3,908	3,908	3,908	3,732
R^2 / within R^2	0.011	0.118	0.018	0.028
Illegal/ Street drugs				
In t	0.077***	0.023	0.073**	0.043
	(0.028)	(0.028)	(0.034)	(0.037)
In t-1	0.093***	0.048*	0.080**	0.037
	(0.027)	(0.026)	(0.033)	(0.034)
N	3,913	3,913	3,913	3,732
R^2 / within R^2	0.005	0.117	0.020	0.028
Tobacco - daily use				
In t	0.044*	0.040*	-0.024	-0.027
	(0.024)	(0.024)	(0.033)	(0.035)
In t-1	0.014	0.007	-0.017	-0.031
	(0.024)	(0.024)	(0.033)	(0.035)
N	3,916	3,916	3,916	3,732
R^2 / within R^2	0.005	0.118	0.016	0.028

Notes: Estimated coefficients and clustered standard errors (at the individual level). The first three models are estimated separately by substance. The first model includes only wave dummies. The second adds the full set of observable time-invariant and time-varying controls described in Table 1 and dummies for missing controls. The third model replaces time-invariant controls with individual fixed effects. The fourth model is identical to the third model, but with all substance use variables entered together.

Table 6: Determinants of substance use at time t

	No controls	Observable controls	Fixed effects
		Alcohol - 21+ standard drinks/wk	
Homeless in t	0.048***	0.030**	0.013
	(0.014)	(0.014)	(0.014)
Homeless in t-1	-0.012	-0.015	-0.024*
	(0.014)	(0.013)	(0.014)
N	3,832	3,832	3,832
R^2 / within R^2	0.004	0.111	0.018
		Cannabis	
Homeless in t	0.081***	0.030*	0.003
	(0.017)	(0.016)	(0.016)
Homeless in t-1	0.056***	0.033**	-0.012
	(0.017)	(0.015)	(0.016)
N	3,899	3,899	3,899
R^2 / within R^2	0.018	0.266	0.040
		Illegal / Street drugs	
Homeless in t	0.042***	0.019	0.018
	(0.012)	(0.012)	(0.013)
Homeless in t-1	0.007	-0.006	-0.005
	(0.012)	(0.011)	(0.013)
N	3,899	3,899	3,899
R^2 / within R^2	0.010	0.153	0.037
		Tobacco - daily use	
Homeless in t	0.019	0.014	-0.009
	(0.017)	(0.017)	(0.014)
Homeless in t-1	0.059***	0.058***	-0.003
	(0.017)	(0.017)	(0.013)
N	3,904	3,904	3,904
R^2 / within R^2	0.006	0.118	0.008

Notes: Estimated coefficients and clustered standard errors (at the individual level), estimated separately for each substance-use behaviour. In each case the first model includes only wave dummies, the second adds the full set of observable time-invariant and time-varying controls described in Table 1 and dummies for missing controls, and the third replaces time-invariant controls with individual fixed effects.

Table 7: Determinants of primary homelessness and degree of homelessness at time t, fixed effects model

	Primary homeless	Degree of Homelessness
Alcohol - risky drinking		
In t	0.015	0.089**
	(0.013)	(0.035)
In t-1	0.036**	0.143***
	(0.014)	(0.037)
Cannabis		
In t	-0.011	0.001
	(0.013)	(0.033)
In t-1	0.001	0.040
	(0.012)	(0.030)
Illegal/ Street drugs		
In t	0.034**	0.074*
	(0.016)	(0.042)
In t-1	-0.031**	0.006
	(0.015)	(0.040)
N	3,749	3,729
R^2 / within R^2	0.025	0.030
Tobacco - daily use		
In t	0.002	-0.022
	(0.011)	(0.038)
In t-1	-0.006	-0.037
	(0.012)	(0.038)

Notes: Estimated coefficients and clustered standard errors (at the individual level) for (1) with individual fixed effects, time-varying controls, and all substance use variables entered together. "Primary homeless" is a dummy equal to 1 if the individual has been primary homeless since the last wave and 0 otherwise; "degree of homelessness is an ordered categorical variable equal to 2 for those having experienced primary homelessness since the last wave, equal to 1 for those having experienced on secondary or tertiary homelessness since the last wave, and equal to 0 otherwise.

Table 8: Determinants of substance use at time t, by type of homelessness, fixed effects model

	Alcohol - risky drinking	Cannabis	Illegal/ Street drugs	Tobacco - daily use
Primary homelessness				
Homeless in t	0.046	0.017	0.100***	0.022
	(0.036)	(0.042)	(0.035)	(0.030)
Homeless in t-1	-0.021	0.024	-0.013	0.035
	(0.033)	(0.031)	(0.031)	(0.025)
Secondary homelessness				
Homeless in t	0.028*	-0.001	0.033**	-0.009
	(0.015)	(0.018)	(0.013)	(0.014)
Homeless in t-1	-0.021	0.015	-0.006	-0.009
	(0.015)	(0.016)	(0.013)	(0.013)
Tertiary homelessness				
Homeless in t	-0.044*	0.028	-0.031	0.007
	(0.025)	(0.029)	(0.021)	(0.020)
Homeless in t-1	0.019	0.025	0.031	0.003
	(0.022)	(0.023)	(0.021)	(0.018)
N	3,809	3,872	3,872	3,877
R^2 / within R^2	0.021	0.041	0.045	0.009

Notes: Estimated coefficients and clustered standard errors (at the individual level) for estimates of (2) with individual fixed effects and time varying controls.

Figure 1: Substance use among users (average over the four waves)